

SEQUENCE LISTING

The following Sequence Listing is submitted pursuant to 37 CFR §1.821. A copy in computer readable form is also submitted herewith. The paper and computer readable forms of this Sequence Listing are the same.

(1) GENERAL INFORMATION:

(i) APPLICANT: Tripp, Cynthia Ann
Frank, Glenn R.
Grieve, Robert B.

(ii) TITLE OF INVENTION: NOVEL PARASITE ASTACIN
METALLOENDOPEPTIDASE PROTEINS

(iii) NUMBER OF SEQUENCES: 36

(iv) CORRESPONDENCE ADDRESS:

(A) ADDRESSEE: SHERIDAN ROSS P.C.
(B) STREET: 1700 LINCOLN ST., SUITE 3500
(C) CITY: DENVER
(D) STATE: CO
(E) COUNTRY: USA
(F) ZIP: 80203

(v) COMPUTER READABLE FORM:

(A) MEDIUM TYPE: Floppy disk
(B) COMPUTER: IBM PC compatible
(C) OPERATING SYSTEM: PC-DOS/MS-DOS
(D) SOFTWARE: PatentIn Release #1.0, Version #1.30

(vi) CURRENT APPLICATION DATA:

(A) APPLICATION NUMBER:
(B) FILING DATE:
(C) CLASSIFICATION:

(viii) ATTORNEY/AGENT INFORMATION:

(A) NAME: Connell, Gary J.
(B) REGISTRATION NUMBER: 32,020
(C) REFERENCE/DOCKET NUMBER: 2618-21-1-C1

(ix) TELECOMMUNICATION INFORMATION:

(A) TELEPHONE: (303) 863-9700
(B) TELEFAX: (303) 863-0223

T.0130

09003574-010698

(2) INFORMATION FOR SEQ ID NO:1:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 1299 base pairs

(B) TYPE: nucleic acid

(C) STRANDEDNESS: single

(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:1:

TTTTTTTTTT TTTTTTTTGT TTCATTGTTC AGTCAGTGGG AAATTATCGA ACGCAGAAAAG 60
 10 CATCACGAAA TACGTTAGAT CACATCAAAC AACTTATCAC CTTGAACGTA CAAAGAGAGA 120
 TTGGAAACAT AGATGATAAG ACATTAGCTG ATGAAATAGT ATTACAACGA CGGGATCCTG 180
 AGGCAAAATG GCATCATAAT GAACTATTCA TTAATGATCC AGATGCATAC TATCAAGGCG 240
 ATGTCGATTT GTCGGAAAAA CAAGCCGAAA TTCTAAGCGA ACATTTTAAA AATGAAATTG 300
 CTTTAACAGA GAAAGACGAC ACAATAATAC GCGGAAAAAA GAGCATTGGT CGTGAACCAT 360
 15 TTTACGTAAG ATGGAATCAT AAACGTCCCA TTAGCTATGA ATTTGCGGAA AGTATTCCAT 420
 TAGAAACACG TAGAAAAATT CGTTCAGCAA TAGCAATGTG GGAAGAACGA ACATGCATAC 480
 GATTCCAAGA AAATGGCCCA AATGTAGATC GAATTGAATT TTACGACGGT GCGGGTTGTT 540
 CAAGTTTGT CGGCCGAACA GGAGGGAATT TCAATTTCAA CACCAGGATG TGATATTATT 600
 GGTATTATAT CACATGAAAT TGGTCATACT TTAGGAATAT TTCATGAGCA AGCACGTCGT 660
 20 GATCAAAAAA ATCATATTTT TATTAATTAC AACAATATTC CATCAAGCCG TTGGAACAAT 720
 TTTTTTCCAT TATCAGAATA TGAAGCTGAT ATGTTTAATT TACCTTATGA TACAGGATCA 780
 GTAATGCACT ATGGTTCATA CGGATTGCA AGAAATCCGT ATGAACCAAC TATTACAACA 840
 CGTGATAAAT TTCAACAGTA CACAATTGGG CAACGTGAAG GGCCATCATT TCTGGATTAT 900
 GCATCTGTTA AGCTTTATCT ACAAACGCAT TAATGATATT GTTATCAAAT GGATGATAAT 960
 25 TTCAATAAGT ATAAACAGCG CTTATCGTTG TACAGAACAA TGTGCTGATA TGCCTGCGA 1020
 TCATAATGGT TATCCGGATC CTAATAATTG CGCGAAATGC TTGTGTCCAG ATGGTTTTGC 1080
 TGGTCGTACC TGTCAATTTG TTCAATATAC ATCTTGCGGA GCTCTCATT AAGTAAGTAT 1140
 TGTCTTTTGA CCTCTTCTCT GACTAAAATA TAAGTTAAGC ATATGTATCT TCCGCTCTAAT 1200
 GATTTTCTTG ATTTTGATTT GTTCAATGCT CTTCTTGATA ATAATATAAA AATTTTGGAA 1260
 30 AATAAAGTTA ACTTTTGGTC AAAAAAAAAA AAAAAAAAAA 1299

(2) INFORMATION FOR SEQ ID NO:2:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 2126 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: single
 (D) TOPOLOGY: linear

5

(ii) MOLECULE TYPE: cDNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:2:

GAAAGCATCA CGAATACGT TAGATCACAT CAAACAACTT ATCACCTTGA ACGTACAAAG 60
 10 AGAGATTGGA AACATAGATG ATAAGACATT AGCTGATGAA ATAGTATTAC AACGACGGGA 120
 TCCTGAGGCA AAATGGCATC ATAATGAACT ATTCATTAAT GATCCAGATG CATACTATCA 180
 AGGCGATGTC GATTTGTCGG AAAACAAGC CGAAATTCTA AGCGAACATT TTAAAAATGA 240
 AATTGCTTTA ACAGAGAAAG ACGACACAAT AATACGGCGA AAAAAGAGCA TTGGTCGTGA 300
 ACCATTTTAC GTAAGATGGA ATCATAAACG TCCCATTAGC TATGAATTTG CGGAAAAGTAT 360
 15 TCCATTAGAA ACACGTAGAA AAATTCGTTC AGCAATAGCA ATGTGGGAAG AACGAACATG 420
 CATACGATTC CAAGAAAATG GCCCAAATGT AGATCGAATT GAATTTTACG ACGGTGGCGG 480
 TTGTTCAAGT TTTGTCGGCC GAACAGGAGG GAATTTCAAT TTCAACACCA GGATGTGATA 540
 TTATTGGTAT TATATCACAT GAAATTGGTC ATACTTTAGG AATATTTTAT GAGCAAGCAC 600
 GTCGTGATCA AAAAATCAT ATTTTTATTA ATTACAACAA TATTCCATCA AGCCGTGGA 660
 20 ACAATTTTTT TCCATTATCA GAATATGAAG CTGATATGTT TAATTTACCT TATGATACAG 720
 GATCAGTAAT GCACTATGGT TCATACGGAT TTGCAAGAAA TCCGTATGAA CCAACTATTA 780
 CAACACGTGA TAAATTTCAA CAGTACACAA TTGGGCAACG TGAAGGGCCA TCATTTCTGG 840
 ATTATGCATC TGATAAACAG CGCTTATCGT TGTACAGAAC AATGTGCTGA TATGCACTGC 900
 GATCATAATG GTTATCCGGA TCCTAATAAT TGCGCGAAAT GCTTGTGTCC AGATGGTTTT 960
 25 GCTGGTCGTA CCTGTCAATT TGTCAATAT ACATCTTGCG GAGCTCTCAT TAAGGCGAGG 1020
 AAAATGCCTG TTACGATTC GAGCCCAAAT TATCCAAACT TCTTCAATGT TGGTGATCAA 1080
 TGTATTTGGT TGCTTACAGC TCCACGCGTG ATTCGTAAAT TTGCAGTTTG TTGAACAATT 1140
 TCAATTACAA TGTGAAGATA CGTGTGATAA ATCCTATGTA GAAGTGAAAG CTGACGCTGA 1200
 TTTTCGACCT ACTGGATATC GATTTTGTTG TTCGCGAGTG CCACGTCATA TTTTCAATC 1260
 30 TGCGACAAAC GAGATGGTAG TAATATTTTCG CGGTTTTGGT GATGCGGGAA ATGGCTTTAA 1320

AGCTAAAATT TGGTCAAACG TAGATGATGA TATAGCTAAT ACAATTGTAA CAACTGAAAT 1380
 GGCAAAAATT TCGGAAAAAA TACCGAAGCT AACAGTTCCA ATAGTTAAAA CTATTACCAC 1440
 TCCTACAATA ACAACTACTA CTGCTTTCAT GATATCACCC AAGAAAGGCA ATGTCACCGC 1500
 CACGAGAGTT GCTATCACTA CTACGCCGAC TACTACAATT ACTACGACTA TTGCCGGTAC 1560
 5 GTACCAATCA CCGTAACTAA TAATACTACA CCTGTAGTAA GTGAAACTTT ACCATCATTG 1620
 CCAGTCAAGA TTCGAAACAA AATAGGTGCA TGCGAATGTG GTGAATGGAC AGAATGGACA 1680
 GGTCCATGCT CTCAAGAATG TGGCGGTTGC GGAAAACGTC TTCGAACACG TCAGTGTTCA 1740
 TCAGATACGG AATGTAGAAC AGAAGAAAAA CGTGCGTGTG CTTTAAAGTT TGCCCATACG 1800
 GGACTAATTT CCTTATCAAT AATGGAGAGT TTCATATACT TTGGAAGGGC TGCTGTGTTG 1860
 10 GTCTATTCCG ATCGGGAGAT ATGTGTTTCA CACTTGATGA TAACGAGAAT CCATTTCTGA 1920
 AATTTCTAGA ATCACTGTTG AACATGCAAG ATTCTCGAAA AAACGATAAT TTGCCTGACT 1980
 CGAAAAAGAA GTGATTGAAT GATTCGATAA TATTGATTAA TAAAACGGGT TGTATTCTCG 2040
 TCATAGAGTA TCCGTTGATG TTTTATCCA AAAAATTCTC TTGCTTTTAA TTATTGTGAA 2100
 TAAAACTTTT GTTTACCCAA AAAAAA 2126

15 (2) INFORMATION FOR SEQ ID NO:3:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 191 amino acids

(B) TYPE: amino acid

(C) STRANDEDNESS:

20 (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:3:

Cys Phe Ile Val Gln Ser Val Glu Asn Tyr Arg Thr Gln Lys Ala Ser
 1 5 10 15

25 Arg Asn Thr Leu Asp His Ile Lys Gln Leu Ile Thr Leu Asn Val Gln
 20 25 30

Arg Glu Ile Gly Asn Ile Asp Asp Lys Thr Leu Ala Asp Glu Ile Val
 35 40 45

30 Leu Gln Arg Arg Asp Pro Glu Ala Lys Trp His His Asn Glu Leu Phe
 50 55 60

Ile Asn Asp Pro Asp Ala Tyr Tyr Gln Gly Asp Val Asp Leu Ser Glu
 65 70 75 80

Lys Gln Ala Glu Ile Leu Ser Glu His Phe Lys Asn Glu Ile Ala Leu
 85 90 95
 Thr Glu Lys Asp Asp Thr Ile Ile Arg Arg Lys Lys Ser Ile Gly Arg
 100 105 110
 5 Glu Pro Phe Tyr Val Arg Trp Asn His Lys Arg Pro Ile Ser Tyr Glu
 115 120 125
 Phe Ala Glu Ser Ile Pro Leu Glu Thr Arg Arg Lys Ile Arg Ser Ala
 130 135 140
 10 Ile Ala Met Trp Glu Glu Arg Thr Cys Ile Arg Phe Gln Glu Asn Gly
 145 150 155 160
 Pro Asn Val Asp Arg Ile Glu Phe Tyr Asp Gly Gly Gly Cys Ser Ser
 165 170 175
 Phe Val Gly Arg Thr Gly Gly Asn Phe Asn Phe Asn Thr Arg Met
 180 185 190

15 (2) INFORMATION FOR SEQ ID NO:4:

- (i) SEQUENCE CHARACTERISTICS:
 (A) LENGTH: 141 amino acids
 (B) TYPE: amino acid
 (C) STRANDEDNESS:
 20 (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:4:

Ile Glu Leu Asn Phe Thr Thr Val Ala Val Val Gln Val Leu Ser Ala
 1 5 10 15
 25 Glu Gln Glu Gly Ile Ser Ile Ser Thr Pro Gly Cys Asp Ile Ile Gly
 20 25 30
 Ile Ile Ser His Glu Ile Gly His Thr Leu Gly Ile Phe His Glu Gln
 35 40 45
 30 Ala Arg Arg Asp Gln Lys Asn His Ile Phe Ile Asn Tyr Asn Asn Ile
 50 55 60
 Pro Ser Ser Arg Trp Asn Asn Phe Phe Pro Leu Ser Glu Tyr Glu Ala
 65 70 75 80
 Asp Met Phe Asn Leu Pro Tyr Asp Thr Gly Ser Val Met His Tyr Gly
 85 90 95
 35 Ser Tyr Gly Phe Ala Arg Asn Pro Tyr Glu Pro Thr Ile Thr Thr Arg
 100 105 110
 Asp Lys Phe Gln Gln Tyr Thr Ile Gly Gln Arg Glu Gly Pro Ser Phe
 115 120 125

Leu Asp Tyr Ala Ser Val Lys Leu Tyr Leu Gln Thr His
 130 135 140

(2) INFORMATION FOR SEQ ID NO:5:

- 5 (i) SEQUENCE CHARACTERISTICS:
 (A) LENGTH: 121 amino acids
 (B) TYPE: amino acid
 (C) STRANDEDNESS:
 (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

10 (xi) SEQUENCE DESCRIPTION: SEQ ID NO:5:

Cys Thr Met Val His Thr Asp Leu Gln Glu Ile Arg Met Asn Gln Leu
 1 5 10 15
 Leu Gln His Val Ile Asn Phe Asn Ser Thr Gln Leu Gly Asn Val Lys
 20 25 30
 Gly His His Phe Trp Ile Met His Leu Leu Ser Phe Ile Tyr Lys Arg
 35 40 45
 Ile Asn Asp Ile Val Ile Lys Trp Met Ile Ile Ser Ile Ser Ile Asn
 50 55 60
 Ser Ala Tyr Arg Cys Thr Glu Gln Cys Ala Asp Met His Cys Asp His
 65 70 75 80
 Asn Gly Tyr Pro Asp Pro Asn Asn Cys Ala Lys Cys Leu Cys Pro Asp
 85 90 95
 Gly Phe Ala Gly Arg Thr Cys Gln Phe Val Gln Tyr Thr Ser Cys Gly
 100 105 110
 Ala Leu Ile Lys Val Ser Ile Val Phe
 115 120

(2) INFORMATION FOR SEQ ID NO:6:

- 30 (i) SEQUENCE CHARACTERISTICS:
 (A) LENGTH: 178 amino acids
 (B) TYPE: amino acid
 (C) STRANDEDNESS:
 (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:6:

35 Lys Ala Ser Arg Asn Thr Leu Asp His Ile Lys Gln Leu Ile Thr Leu
 1 5 10 15
 Asn Val Gln Arg Glu Ile Gly Asn Ile Asp Asp Lys Thr Leu Ala Asp
 20 25 30

Glu Ile Val Leu Gln Arg Arg Asp Pro Glu Ala Lys Trp His His Asn
 35 40 45
 Glu Leu Phe Ile Asn Asp Pro Asp Ala Tyr Tyr Gln Gly Asp Val Asp
 50 55 60
 5 Leu Ser Glu Lys Gln Ala Glu Ile Leu Ser Glu His Phe Lys Asn Glu
 65 70 75 80
 Ile Ala Leu Thr Glu Lys Asp Asp Thr Ile Ile Arg Arg Lys Lys Ser
 85 90 95
 10 Ile Gly Arg Glu Pro Phe Tyr Val Arg Trp Asn His Lys Arg Pro Ile
 100 105 110
 Ser Tyr Glu Phe Ala Glu Ser Ile Pro Leu Glu Thr Arg Arg Lys Ile
 115 120 125
 Arg Ser Ala Ile Ala Met Trp Glu Glu Arg Thr Cys Ile Arg Phe Gln
 130 135 140
 15 Glu Asn Gly Pro Asn Val Asp Arg Ile Glu Phe Tyr Asp Gly Gly Gly
 145 150 155 160
 Cys Ser Ser Phe Val Gly Arg Thr Gly Gly Asn Phe Asn Phe Asn Thr
 165 170 175
 20 Arg Met

(2) INFORMATION FOR SEQ ID NO:7:

- (i) SEQUENCE CHARACTERISTICS:
 (A) LENGTH: 145 amino acids
 (B) TYPE: amino acid
 25 (C) STRANDEDNESS:
 (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:7:

30 Ile Glu Leu Asn Phe Thr Thr Val Ala Val Val Gln Val Leu Ser Ala
 1 5 10 15
 Glu Gln Glu Gly Ile Ser Ile Ser Thr Pro Gly Cys Asp Ile Ile Gly
 20 25 30
 Ile Ile Ser His Glu Ile Gly His Thr Leu Gly Ile Phe His Glu Gln
 35 40 45
 35 Ala Arg Arg Asp Gln Lys Asn His Ile Phe Ile Asn Tyr Asn Asn Ile
 50 55 60
 Pro Ser Ser Arg Trp Asn Asn Phe Phe Pro Leu Ser Glu Tyr Glu Ala
 65 70 75 80

Asp Met Phe Asn Leu Pro Tyr Asp Thr Gly Ser Val Met His Tyr Gly
 85 90 95
 Ser Tyr Gly Phe Ala Arg Asn Pro Tyr Glu Pro Thr Ile Thr Thr Arg
 100 105 110
 5 Asp Lys Phe Gln Gln Tyr Thr Ile Gly Gln Arg Glu Gly Pro Ser Phe
 115 120 125
 Leu Asp Tyr Ala Ser Asp Lys Gln Arg Leu Ser Leu Tyr Arg Thr Met
 130 135 140
 10 Cys
 145

(2) INFORMATION FOR SEQ ID NO:8:

(i) SEQUENCE CHARACTERISTICS:
 (A) LENGTH: 134 amino acids
 (B) TYPE: amino acid
 15 (C) STRANDEDNESS:
 (D) TOPOLOGY: linear
 (ii) MOLECULE TYPE: protein
 (xi) SEQUENCE DESCRIPTION: SEQ ID NO:8:
 20 Cys Thr Met Val His Thr Asp Leu Gln Glu Ile Arg Met Asn Gln Leu
 1 5 10 15
 Leu Gln His Val Ile Asn Phe Asn Ser Thr Gln Leu Gly Asn Val Lys
 20 25 30
 Gly His His Phe Trp Ile Met His Leu Ile Asn Ser Ala Tyr Arg Cys
 35 40 45
 25 Thr Glu Gln Cys Ala Asp Met His Cys Asp His Asn Gly Tyr Pro Asp
 50 55 60
 Pro Asn Asn Cys Ala Lys Cys Leu Cys Pro Asp Gly Phe Ala Gly Arg
 65 70 75 80
 30 Thr Cys Gln Phe Val Gln Tyr Thr Ser Cys Gly Ala Leu Ile Lys Ala
 85 90 95
 Arg Lys Met Pro Val Thr Ile Ser Ser Pro Asn Tyr Pro Asn Phe Phe
 100 105 110
 Asn Val Gly Asp Gln Cys Ile Trp Leu Leu Thr Ala Pro Arg Val Ile
 115 120 125
 35 Arg Lys Phe Ala Val Cys
 130

(2) INFORMATION FOR SEQ ID NO:9:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 154 amino acids
 (B) TYPE: amino acid
 (C) STRANDEDNESS:
 (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:9:

10 Phe Val Asn Leu Gln Phe Val Glu Gln Phe Gln Leu Gln Cys Glu Asp
 1 5 10 15
 Thr Cys Asp Lys Ser Tyr Val Glu Val Lys Ala Asp Ala Asp Phe Arg
 20 25 30
 Pro Thr Gly Tyr Arg Phe Cys Cys Ser Arg Val Pro Arg His Ile Phe
 35 40 45
 15 Gln Ser Ala Thr Asn Glu Met Val Val Ile Phe Arg Gly Phe Gly Asp
 50 55 60
 Ala Gly Asn Gly Phe Lys Ala Lys Ile Trp Ser Asn Val Asp Asp Asp
 65 70 75 80
 20 Ile Ala Asn Thr Ile Val Thr Thr Glu Met Ala Lys Ile Ser Glu Lys
 85 90 95
 Ile Pro Lys Leu Thr Val Pro Ile Val Lys Thr Ile Thr Thr Pro Thr
 100 105 110
 Ile Thr Thr Thr Thr Ala Phe Met Ile Ser Pro Lys Lys Gly Asn Val
 115 120 125
 25 Thr Ala Thr Arg Asx Ala Ile Thr Thr Thr Pro Thr Thr Thr Ile Thr
 130 135 140
 Thr Thr Ile Ala Gly Thr Tyr Gln Ser Pro
 145 150

(2) INFORMATION FOR SEQ ID NO:10:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 163 amino acids
 (B) TYPE: amino acid
 (C) STRANDEDNESS:
 (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:10:

Asn Tyr Tyr His Ser Tyr Asn Asn Asn Tyr Tyr Cys Phe His Asp Ile
 1 5 10 15

Thr Gln Glu Arg Gln Cys His Arg His Glu Ser Cys Tyr His Tyr Tyr
 20 25 30
 Ala Asp Tyr Tyr Asn Tyr Tyr Asp Tyr Cys Arg Tyr Val Pro Ile Thr
 35 40 45
 5 Val Thr Asn Asn Thr Thr Pro Val Val Ser Glu Thr Leu Pro Ser Leu
 50 55 60
 Pro Val Lys Ile Arg Asn Lys Ile Gly Ala Cys Glu Cys Gly Glu Trp
 65 70 75 80
 10 Thr Glu Trp Thr Gly Pro Cys Ser Gln Glu Cys Gly Gly Cys Gly Lys
 85 90 95
 Arg Leu Arg Thr Arg Gln Cys Ser Ser Asp Thr Glu Cys Arg Thr Glu
 100 105 110
 Glu Lys Arg Ala Cys Ala Phe Lys Phe Ala His Thr Gly Leu Ile Ser
 115 120 125
 15 Leu Ser Ile Met Glu Ser Phe Ile Tyr Phe Gly Arg Ala Ala Val Leu
 130 135 140
 Val Tyr Ser Asp Arg Glu Ile Cys Val Gln His Leu Met Ile Thr Arg
 145 150 155 160
 20 Ile His Phe

(2) INFORMATION FOR SEQ ID NO:11:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 638 amino acids
 (B) TYPE: amino acid
 (C) STRANDEDNESS:
 (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:11:

Lys Ala Ser Arg Asn Thr Leu Asp His Ile Lys Gln Leu Ile Thr Leu
 1 5 10 15
 Asn Val Gln Arg Glu Ile Gly Asn Ile Asp Asp Lys Thr Leu Ala Asp
 20 25 30
 Glu Ile Val Leu Gln Arg Arg Asp Pro Glu Ala Lys Trp His His Asn
 35 40 45
 35 Glu Leu Phe Ile Asn Asp Pro Asp Ala Tyr Tyr Gln Gly Asp Val Asp
 50 55 60

Leu Ser Glu Lys Gln Ala Glu Ile Leu Ser Glu His Phe Lys Leu Asn
 65 70 75 80
 Glu Ile Ala Leu Thr Glu Lys Asp Asp Thr Ile Ile Arg Arg Lys Lys
 85 90 95
 5 Ser Ile Gly Arg Glu Pro Phe Tyr Val Arg Trp Asn His Lys Arg Pro
 100 105 110
 Ile Ser Tyr Glu Phe Ala Glu Ser Ile Pro Leu Glu Thr Arg Arg Lys
 115 120 125
 10 Ile Arg Ser Ala Ile Ala Met Trp Glu Glu Arg Thr Cys Ile Arg Phe
 130 135 140
 Gln Glu Asn Gly Pro Asn Val Asp Arg Ile Glu Phe Tyr Asp Gly Gly
 145 150 155 160
 Gly Cys Ser Ser Phe Val Gly Arg Gln Glu Gly Ile Ser Ile Ser Thr
 165 170 175
 15 Pro Gly Cys Asp Ile Ile Gly Ile Ile Ser His Glu Ile Gly His Thr
 180 185 190
 Leu Gly Ile Phe His Glu Gln Ala Arg Arg Asp Gln Lys Asn His Ile
 195 200 205
 20 Phe Ile Asn Tyr Asn Asn Ile Pro Ser Ser Arg Trp Asn Asn Phe Phe
 210 215 220
 Pro Leu Ser Glu Tyr Glu Ala Asp Met Phe Asn Leu Pro Tyr Asp Thr
 225 230 235 240
 Gly Ser Val Met His Tyr Gly Ser Tyr Gly Phe Ala Arg Asn Pro Tyr
 245 250 255
 25 Glu Pro Thr Ile Thr Thr Arg Asp Lys Phe Gln Gln Tyr Thr Ile Gly
 260 265 270
 Gln Arg Glu Gly Pro Ser Phe Leu Asp Met His Leu Ile Asn Ser Ala
 275 280 285
 30 Tyr Arg Cys Thr Glu Gln Cys Ala Asp Met His Cys Asp His Asn Gly
 290 295 300
 Tyr Pro Asp Pro Asn Asn Cys Ala Lys Cys Leu Cys Pro Asp Gly Phe
 305 310 315 320
 Ala Gly Arg Thr Cys Gln Phe Val Gln Tyr Thr Ser Cys Gly Ala Leu
 325 330 335
 35 Ile Lys Ala Arg Lys Met Pro Val Thr Ile Ser Ser Pro Asn Tyr Pro
 340 345 350
 Asn Phe Phe Asn Tyr Gly Asp Gln Cys Ile Trp Leu Leu Thr Ala Pro
 355 360 365

09003574-010698

Arg Val Phe Val Asn Leu Gln Phe Val Glu Gln Phe Gln Leu Gln Cys
 370 375 380
 Glu Asp Thr Cys Asp Lys Ser Tyr Val Glu Val Lys Ala Asp Ala Asp
 385 390 395 400
 5 Phe Arg Pro Thr Gly Tyr Arg Phe Cys Cys Ser Arg Val Pro Arg His
 405 410 415
 Ile Phe Gln Ser Ala Thr Asn Glu Met Val Val Ile Phe Arg Gly Phe
 420 425 430
 10 Gly Asp Ala Gly Asn Gly Phe Lys Ala Lys Ile Trp Ser Asn Val Asp
 435 440 445
 Asp Asp Ile Ala Asn Thr Ile Val Thr Thr Glu Met Ala Lys Ile Ser
 450 455 460
 Glu Lys Ile Pro Lys Leu Thr Val Pro Ile Val Lys Thr Ile Thr Thr
 465 470 475 480
 15 Pro Thr Ile Thr Thr Thr Thr Ala Phe Met Ile Ser Pro Lys Lys Gly
 485 490 495
 Asn Val Thr Ala Thr Arg Val Ala Ile Thr Thr Thr Pro Thr Thr Thr
 500 505 510
 20 Ile Thr Thr Thr Ile Ala Gly Thr Tyr Gln Ser Val Thr Asn Asn Thr
 515 520 525
 Thr Pro Val Val Ser Glu Thr Leu Pro Ser Leu Pro Val Lys Ile Arg
 530 535 540
 Asn Lys Ile Gly Ala Cys Glu Cys Gly Glu Trp Thr Glu Trp Thr Gly
 545 550 555 560
 25 Pro Cys Ser Gln Glu Cys Gly Gly Cys Gly Lys Arg Leu Arg Thr Arg
 565 570 575
 Gln Cys Ser Ser Asp Thr Glu Cys Arg Thr Glu Glu Lys Arg Ala Cys
 580 585 590
 30 Ala Phe Lys Phe Ala His Thr Gly Leu Ile Ser Leu Ser Ile Met Glu
 595 600 605
 Ser Phe Ile Tyr Phe Gly Arg Ala Ala Val Leu Val Tyr Ser Asp Arg
 610 615 620
 Glu Ile Cys Val Gln His Leu Met Ile Thr Arg Ile His Phe
 625 630 635

(2) INFORMATION FOR SEQ ID NO:12:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 143 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: single
 (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(ix) FEATURE:

- (A) NAME/KEY: CDS
 (B) LOCATION: 1..143

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:12:

GGA TCC TGT GGT TCA TGT TGG GCT TTT TCT GTT ACT GGC AAT ATT GCA 48
 Gly Ser Cys Gly Ser Cys Trp Ala Phe Ser Val Thr Gly Asn Ile Ala
 1 5 10 15

AGT CTC TGG GCT ATT AAA ACA GGT GAT TTG ATA TCG CTT TCC GAG CAA 96
 Ser Leu Trp Ala Ile Lys Thr Gly Asp Leu Ile Ser Leu Ser Glu Gln
 20 25 30

GAA TTG ATT GAT TGT GAT GTG GTT GAT GAG GGC TGC AAC GGC GGC TA 143
 Glu Leu Ile Asp Cys Asp Val Val Asp Glu Gly Cys Asn Gly Gly
 35 40 45

(2) INFORMATION FOR SEQ ID NO:13:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 47 amino acids
 (B) TYPE: amino acid
 (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:13:

Gly Ser Cys Gly Ser Cys Trp Ala Phe Ser Val Thr Gly Asn Ile Ala
 1 5 10 15

Ser Leu Trp Ala Ile Lys Thr Gly Asp Leu Ile Ser Leu Ser Glu Gln
 20 25 30

Glu Leu Ile Asp Cys Asp Val Val Asp Glu Gly Cys Asn Gly Gly
 35 40 45

(2) INFORMATION FOR SEQ ID NO:14:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 18 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: single
 (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)
 (ix) FEATURE:
 (A) NAME/KEY: misc_feature
 (B) LOCATION: 12
 5 (D) OTHER INFORMATION: /label= INOSINE
 (ix) FEATURE:
 (A) NAME/KEY: misc_feature
 (B) LOCATION: 15
 (D) OTHER INFORMATION: /label= INOSINE
 10 (xi) SEQUENCE DESCRIPTION: SEQ ID NO:14:
 ACWCATGAAA TNGSNCAT 18
 (2) INFORMATION FOR SEQ ID NO:15:
 (i) SEQUENCE CHARACTERISTICS:
 (A) LENGTH: 17 base pairs
 15 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: single
 (D) TOPOLOGY: linear
 (ii) MOLECULE TYPE: DNA (genomic)
 (xi) SEQUENCE DESCRIPTION: SEQ ID NO:15:
 20 AATACGACTC ACTATAG 17
 (2) INFORMATION FOR SEQ ID NO:16:
 (i) SEQUENCE CHARACTERISTICS:
 (A) LENGTH: 30 base pairs
 25 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: single
 (D) TOPOLOGY: linear
 (ii) MOLECULE TYPE: DNA (genomic)
 (xi) SEQUENCE DESCRIPTION: SEQ ID NO:16:
 TGGTATTATA TCACATGAAA TTGGTCATAC 30
 30 (2) INFORMATION FOR SEQ ID NO:17:
 (i) SEQUENCE CHARACTERISTICS:
 (A) LENGTH: 29 base pairs
 (B) TYPE: nucleic acid
 35 (C) STRANDEDNESS: single
 (D) TOPOLOGY: linear
 (ii) MOLECULE TYPE: DNA (genomic)

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:17:

CCCAATTGTG TACTGTTGAA ATTTATCAC

29

(2) INFORMATION FOR SEQ ID NO:18:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 25 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:18:

CGGGATCCTG TGGWTCATGY TGGGC

25

(2) INFORMATION FOR SEQ ID NO:19:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 20 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:19:

TANCCNCCRT TRCANCCYTC

20

(2) INFORMATION FOR SEQ ID NO:20:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 689 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:20:

TCACATGAAA TTGGTCATAC TTTAGGAATA TTTCATGAGC AAGCACGTCG TGATCAAAAA

60

AATCATATTT TTATTAATTA CAACAATATT CCATCAAGCC GTTGAACAA TTTTTTTCCA

120

TTATCAGAAT ATGAAGCTGA TATGTTTAAT TTACCTTATG ATACAGGATC AGTAATGCAC

180

TATGGTTCAT ACGGATTTGC AAGAAATCCG TATGAACCAA CTATTACAAC ACGTGATAAA

240

TTTCAACAGT ACACAATTGG GCAACGTGAA GGGCCATCAT TTCTGGATTA TGCATCTGTT

300

AAGCTTTATC TACAAACGCA TTAATGATAT TGTTATCAAA TGGATGATAA TTTCAATAAG

360

TATAAACAGC GCTTATCGTT GTACAGAACA ATGTGCTGAT ATGCACTGCG ATCATAATGG 420
 TTATCCGGAT CCTAATAATT GCGCGAAATG CTTGTGTCCA GATGGTTTTG CTGGTCGTAC 480
 CTGTCAATTT GTTCAATATA CATCTTGCGG AGCTCTCATT AAGGTAAGTA TTGTCTTTTG 540
 ACCTCTTCTC TGAATAAAAT ATAAGTTAAG CATATGTATC TTCCGTCTAA TGATTTTCTT 600
 5 GATTTTGATT TGTTCATGC TCTTCTTGAT AATAATATAA AAATTTTGA AAATAAAGTT 660
 AACTTTTGGT CAAAAAAAAA AAAAAAAAAA 689

(2) INFORMATION FOR SEQ ID NO:21:

(i) SEQUENCE CHARACTERISTICS:

- 10 (A) LENGTH: 804 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: single
 (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:21:

15 GATCCTGAGG CAAATGGCA TCATAATGAA CTATTCATTA ATGATCCAGA TGCATACTAT 60
 CAAGGCGATG TCGATTTGTC GGAACAAACAA GCCGAAATTC TAAGCGAACA TTTTAAAAAT 120
 GAAATTGCTT TAACAGAGAA AGACGACACA ATAATACGGC GAAAAAGAG CATTGGTCGT 180
 GAACCATTTT ACGTAAGATG GAATCATAAA CGTCCCATTA GCTATGAATT TGCGGAAAGT 240
 ATTCATTAG AAACACGTAG AAAAATTCGT TCAGCAATAG CAATGTGGGA AGAACGAACA 300
 20 TGCATACGAT TCCAAGAAAA TGGCCCAAAT GTAGATCGAA TTGAATTTTA CGACGGTGGC 360
 GGTGTTCAA GTTTTGTCTG CCGAACAGGA GGAATTTCA ATTTCAACAC CAGGATGTGA 420
 TATTATTGGT ATTATATCAC ATGAAATTGG TCATACTTTA GGAATATTTT ATGAGCAAGC 480
 ACGTCGTGAT CAAAAAATC ATATTTTAT TAATTACAAC AATATTCCAT CAAGCCGTTG 540
 GAACAATTTT TTTCCATTAT CAGAAATGA AGCTGATATG TTTAATTTAC CTTATGATAC 600
 25 AGGATCAGTA ATGCACTATG GTTCATACGG ATTTGCAAGA AATCCGTATG AACCAACTAT 660
 TACAACACGT GATAAATTTT AACAGTACAC AATTGGGCAA CGTGAAGGGC CATCATTTCT 720
 GGATTATGCA TCTGATAAAC AGCGCTTATC GTTGACAGA ACAATGTGCT GATATGCACT 780
 GCGATCATAA TGGTTATCCG GATC 804

(2) INFORMATION FOR SEQ ID NO:22:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 271 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: single
 (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:22:

TGGTATTATA TCACATGAAA TTGGTCATAC TTTAGGAATA TTTTCATGAGC AAGCACGTCG 60
 TGATCAAAAA AATCATATTT TTATTAATTA CAACAATATT CCATCAAGCC GTTGGAAACAA 120
 TTTTTTTCCA TTATCAGAAT ATGAAGCTGA TATGTTTAAT TTACCTTATG ATACAGGATC 180
 AGTAATGCAC TATGGTTCAT ACGGATTTGC AAGAAATCCG TATGAACCAA CTATTACAAC 240
 ACGTGATAAA TTTCAACAGT ACACAATTGG G 271

(2) INFORMATION FOR SEQ ID NO:23:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 142 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: single
 (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(ix) FEATURE:

- (A) NAME/KEY: CDS
 (B) LOCATION: 3..140

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:23:

GA TCC TGT GGT TCA TGT TGG GCT TTT TCT GTT ACT GGC AAT ATT GCA 47
 Ser Cys Gly Ser Cys Trp Ala Phe Ser Val Thr Gly Asn Ile Ala
 1 5 10 15
 AGT CTC TGG GCT ATT AAA ACA GGT GAT TTG ATA TCG CTT TCC GAG CAA 95
 Ser Leu Trp Ala Ile Lys Thr Gly Asp Leu Ile Ser Leu Ser Glu Gln
 20 25 30
 GAA TTG ATT GAT TGT GAT GTG GTT GAT GAG GGC TGC AAC GGC GGC 140
 Glu Leu Ile Asp Cys Asp Val Val Asp Glu Gly Cys Asn Gly Gly
 35 40 45
 TA 142

(2) INFORMATION FOR SEQ ID NO:24:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 46 amino acids
 (B) TYPE: amino acid
 (D) TOPOLOGY: linear

5

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:24:

Ser Cys Gly Ser Cys Trp Ala Phe Ser Val Thr Gly Asn Ile Ala Ser
 1 5 10 15
 10 Leu Trp Ala Ile Lys Thr Gly Asp Leu Ile Ser Leu Ser Glu Gln Glu
 20 25 30
 Leu Ile Asp Cys Asp Val Val Asp Glu Gly Cys Asn Gly Gly
 35 40 45

(2) INFORMATION FOR SEQ ID NO:25:

15

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 34 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: single
 (D) TOPOLOGY: linear

20

(ii) MOLECULE TYPE: other nucleic acid

- (A) DESCRIPTION: /desc = "primer"

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:25:

GTCGGATCCG CAGGAGGGAA TTTCAATTC AACA

34

(2) INFORMATION FOR SEQ ID NO:26:

25

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 34 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: single
 (D) TOPOLOGY: linear

30

(ii) MOLECULE TYPE: other nucleic acid

- (A) DESCRIPTION: /desc = "primer"

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:26:

TCAAGATCTA ATCCAGAAAT GATGGCCCTT CACG

34

5 (i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 19 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: other nucleic acid
(A) DESCRIPTION: /desc = "primer"

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:27:

(2) INFORMATION FOR SEQ ID NO:28:

15

- (i) SEQUENCE CHARACTERISTICS:
 - (A) LENGTH: 17 base pairs
 - (B) TYPE: nucleic acid
 - (C) STRANDEDNESS: single
 - (D) TOPOLOGY: linear
- (ii) MOLECULE TYPE: other nucleic acid
 - (A) DESCRIPTION: /desc = "primer"
- (xi) SEQUENCE DESCRIPTION: SEQ ID NO:28:

(2) INFORMATION FOR SEO ID NO:29:

25 (i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 2292 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

30 (ix) FEATURE:
(A) NAME/KEY: misc_feature
(B) LOCATION: 1459
(D) OTHER INFORMATION: /note= "1459:S=C or G;
aa463:Xaa=Alanine or Glycine"

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:29:		
	TAGATTTCTGA TTCGTCTTTG TTAATTCATC TTCGTCAGAT TTATTAGAGA AAAATAAAAA	60
35	TTTTGATCGC AATGAAGCAG GTTATCATCT TTCCTCAGCT TTTCATTTGT TTCATTGTTT	120
	AGTCAGTGGA AAATTATCGA ACGCAGAAAG CATCACGAÄÄ TACGTTAGAT CACATCAAAC	180
	AACTTATCAC CTTGAACGTA CAAAGAGAGA TTGGAAACAT AGATGATAAG ACATTAGCTG	240

5 ATGAAATAGT ATTACAACGA CGGGATCCTG AGGCAAAATG GCATCATAAT GAACTATTCA 300
 TTAATGATCC AGATGCATAC TATCAAGGCG ATGTCGATTT GTCGGAAAAA CAAGCCGAAA 360
 TTCTAAGCGA ACATTTTAAA AATGAAATTG CTTTAAACAGA GAAAGACGAC ACAATAATAC 420
 GCGGAAAAAA GAGCATTGGT CGTGAACCAT TTTACGTAAG ATGGAATCAT AAACGTCCCA 480
 10 TTAGCTATGA ATTTGCGGAA AGTATTCCAT TAGAAAACAG TAGAAAAATT CGTTCAGCAA 540
 TAGCAATGTG GGAAGAACGA ACATGCATAC GATTCCAAGA AAATGGCCCA AATGTTGATC 600
 GAATTGAATT TTACGACGGT GCGGGTTGTT CAAGTTTTGT CCGCCGAACA GGAGGCACGC 660
 AAGGAATTTT AATTTCAACA CCAGGATGTG ATATTATTGG TATTATATCA CATGAAATTG 720
 GTCATACTTT AGGAATATTT CATGAGCAAG CACGTCGTGA TCAAAAAAAT CATATTTTTTA 780
 15 TTAATTACAA CAATATTTCA TCAAGCCGTT GGAACAATTT TTTTCCATTA TCAGAAATAG 840
 AAGCTGATAT GTTTAATTTA CCTTATGATA CAGGATCAGT AATGCACTAT GGTTCATACG 900
 GATTTGCAAG AAATCCGTAT GAACCAACTA TTACAACACG TGATAAATTT CAACAGTACA 960
 CAATTGGGCA ACGTGAAGGG CCATCATTTT TGGATTATGC ATCTATAAAC AGCGCTTATC 1020
 GTTGTACAGA ACAATGTGCT GATATGCACT GCGATCATAA TGGTTATCCG GATCCTAATA 1080
 20 ATTGCGCGAA ATGCTTGTGT CCAGATGGTT TTGCTGGTCG TACCTGTCAA TTTGTTCAAT 1140
 ATACATCTTG CGGAGCTCTC ATTAAGGCGA GGAAAAATGCC TGTACGATT TCGAGCCCAA 1200
 ATTATCCAAA CTTCTTCAAT GTTGGTGATC AATGTATTTG GTTGCTTACA GCTCCACGCG 1260
 GTGGATTTCG AAATTTGCAG TTTGTGTAAC AATTTCAATT ACAATGTGAA GATACGTGTG 1320
 ATAAATCCTA TGTAGAAGTG AAAGCTGACG CTGATTTTCG ACCTACTGGA TATCGATTTT 1380
 25 GTTGTTCGCG AGTGCCACGT CATATTTTTT AATCTGCGAC AAACGAGATG GTAGTAATAT 1440
 TTCGCGGTTT TGGTGATGCG GGAAATGGCT TTAAAGCTAA AATTTGGTCA AACGTAGATG 1500
 ATGATATAGC TAATACAATT GTAACAACTG AAATGGCAAA AATTTCTGGAA AAAATACCGA 1560
 AGCTAACAGT TCCAATAGTT AAACTATTA CCACTCCTAC AATAACAACT ACTACTGCTT 1620
 TCATGATATC ACCCAAGAAA GGCAATGTCA CCGCCACGAG AGTTGCTATC ACTACTACGC 1680
 CGACTACTAC AATTACTACG ACTATTGCCG GTACGGTACC AATCACCGTA ACTAATAATA 1740
 CTACACCTGT AGTAAGTGAA ACTTTACCAT CATTGCCAGT CAAGATTCTGA AACAAAATAG 1800
 GTGCATGCGA ATGTGGTGAA TGGACAGAAT GGACAGGTCC ATGCTCTCAA GAATGTGGCG 1860
 GTTGCGGAAA ACGTCTTCGA ACACGTCAGT GTTCATCAGA TACGGAATGT AGAACAGAAG 1920

5

10

- (ii) MOLECULE TYPE: cDNA

- (A) NAME/KEY: CDS
(B) LOCATION: 1..2076

- 20

CAG TCA GTG GAA AAT TAT CGA ACG CAG AAA GCA TCA CGA AAT ACG TTA 96
Gln Ser Val Glu Asn Tyr Arg Thr Gln Lys Ala Ser Arg Asn Thr Leu
20 25 30

25

GAT CAC ATC AAA CAA CTT ATC ACC TTG AAC GTA CAA AGA GAG ATT GGA 144
Asp His Ile Lys Gln Leu Ile Thr Leu Asn Val Gln Arg Glu Ile Gly
35 40 45

30

AAC ATA GAT GAT AAG ACA TTA GCT GAT GAA ATA GTA TTA CAA CGA CGG 192
Asn Ile Asp Asp Lys Thr Leu Ala Asp Glu Ile Val Leu Gln Arg Arg
50 55 60

GAT CCT GAG GCA AAA TGG CAT CAT AAT GAA CTA TTC ATT AAT GAT CCA 240
Asp Pro Glu Ala Lys Trp His His Asn Glu Leu Phe Ile Asn Asp Pro
65 70 75 80

35

GAT GCA TAC TAT CAA GGC GAT GTC GAT TTG TCG GAA AAA CAA GCC GAA 288
Asp Ala Tyr Tyr Gln Gly Asp Val Asp Leu Ser Glu Lys Gln Ala Glu
85 90 95

152

ATT GCC GGT ACG GTA CCA ATC ACC GTA ACT AAT AAT ACT ACA CCT GTA 1680
 Ile Ala Gly Thr Val Pro Ile Thr Val Thr Asn Asn Thr Thr Pro Val
 545 550 555 560

5 GTA AGT GAA ACT TTA CCA TCA TTG CCA GTC AAG ATT CGA AAC AAA ATA 1728
 Val Ser Glu Thr Leu Pro Ser Leu Pro Val Lys Ile Arg Asn Lys Ile
 565 570 575

GGT GCA TGC GAA TGT GGT GAA TGG ACA GAA TGG ACA GGT CCA TGC TCT 1776
 Gly Ala Cys Glu Cys Gly Glu Trp Thr Glu Trp Thr Gly Pro Cys Ser
 580 585 590

10 CAA GAA TGT GGC GGT TGC GGA AAA CGT CTT CGA ACA CGT CAG TGT TCA 1824
 Gln Glu Cys Gly Gly Cys Gly Lys Arg Leu Arg Thr Arg Gln Cys Ser
 595 600 605

TCA GAT ACG GAA TGT AGA ACA GAA GAA AAA CGT GCG TGT GCT TTT AAA 1872
 Ser Asp Thr Glu Cys Arg Thr Glu Glu Lys Arg Ala Cys Ala Phe Lys
 15 610 615 620

GTT TGC CCA TAC GGG ACT AAT TTC CTT ATC AAT AAT GGA GAG TTT CAT 1920
 Val Cys Pro Tyr Gly Thr Asn Phe Leu Ile Asn Asn Gly Glu Phe His
 625 630 635 640

ATA CTT TGG AAG GGC TGC TGT GTT GGT CTA TTC CGA TCG GGA GAT ATG 1968
 Ile Leu Trp Lys Gly Cys Cys Val Gly Leu Phe Arg Ser Gly Asp Met
 20 645 650 655

TGT TCA GCA CTT GAT GAT AAC GAG AAT CCA TTT CTG AAA TTT CTA GAA 2016
 Cys Ser Ala Leu Asp Asp Asn Glu Asn Pro Phe Leu Lys Phe Leu Glu
 660 665 670

25 TCA CTG TTG AAC ATG CAA GAT TCT CGA AAA AAC GAT AAT TTG CCT GAC 2064
 Ser Leu Leu Asn Met Gln Asp Ser Arg Lys Asn Asp Asn Leu Pro Asp
 675 680 685

TCG AAA AAG AAG 2076
 Ser Lys Lys Lys
 30 690

(2) INFORMATION FOR SEQ ID NO:31:

(i) SEQUENCE CHARACTERISTICS:

- 35 (A) LENGTH: 692 amino acids
 (B) TYPE: amino acid
 (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:31:

Met Lys Gln Val Ile Ile Phe Pro Gln Leu Phe Ile Cys Phe Ile Val
 1 5 10 15

153

Ser Phe Leu Asp Tyr Ala Ser Ile Asn Ser Ala Tyr Arg Cys Thr Glu
 305 310 315 320
 Gln Cys Ala Asp Met His Cys Asp His Asn Gly Tyr Pro Asp Pro Asn
 325 330 335
 5 Asn Cys Ala Lys Cys Leu Cys Pro Asp Gly Phe Ala Gly Arg Thr Cys
 340 345 350
 Gln Phe Val Gln Tyr Thr Ser Cys Gly Ala Leu Ile Lys Ala Arg Lys
 355 360 365
 10 Met Pro Val Thr Ile Ser Ser Pro Asn Tyr Pro Asn Phe Phe Asn Val
 370 375 380
 Gly Asp Gln Cys Ile Trp Leu Leu Thr Ala Pro Arg Gly Gly Phe Val
 385 390 395 400
 Asn Leu Gln Phe Val Glu Gln Phe Gln Leu Gln Cys Glu Asp Thr Cys
 405 410 415
 15 Asp Lys Ser Tyr Val Glu Val Lys Ala Asp Ala Asp Phe Arg Pro Thr
 420 425 430
 Gly Tyr Arg Phe Cys Cys Ser Arg Val Pro Arg His Ile Phe Gln Ser
 435 440 445
 20 Ala Thr Asn Glu Met Val Val Ile Phe Arg Gly Phe Gly Asp Ala Gly
 450 455 460
 Asn Gly Phe Lys Ala Lys Ile Trp Ser Asn Val Asp Asp Asp Ile Ala
 465 470 475 480
 Asn Thr Ile Val Thr Thr Glu Met Ala Lys Ile Ser Glu Lys Ile Pro
 485 490 495
 25 Lys Leu Thr Val Pro Ile Val Lys Thr Ile Thr Thr Pro Thr Ile Thr
 500 505 510
 Thr Thr Thr Ala Phe Met Ile Ser Pro Lys Lys Gly Asn Val Thr Ala
 515 520 525
 30 Thr Arg Val Ala Ile Thr Thr Thr Pro Thr Thr Thr Ile Thr Thr Thr
 530 535 540
 Ile Ala Gly Thr Val Pro Ile Thr Val Thr Asn Asn Thr Thr Pro Val
 545 550 555 560
 Val Ser Glu Thr Leu Pro Ser Leu Pro Val Lys Ile Arg Asn Lys Ile
 565 570 575
 35 Gly Ala Cys Glu Cys Gly Glu Trp Thr Glu Trp Thr Gly Pro Cys Ser
 580 585 590

0900374-010698

TGAAGCTGAT ATGTTTAATT TACCTTATGA TACAGGATCA GTAATGCACT ATGGTTCATA 780
 CGGATTTGCA AGAAATCCGT ATGAACCAAC TATTACAACA CGTGATAAAT TTCAACAGTA 840
 CACAATTGGG CAACGTGAAG GGCCATCATT TCTGGATTAT GCATCTATAA ACAGCGCTTA 900
 TCGTTGTACA GAACAATGTG CTGATATGCA CTGCGATCAT AATGGTTATC CGGATCCTAA 960
 5 TAATTGCGCG AAATGCTTGT GTCCAGATGG TTTTGCTGGT CGTACCTGTC AATTTGTTCA 1020
 ATATACATCT TCGGGAGCTC TCATTAAGGC GAGGAAAATG CCTGTTACGA TTTCGAGCCC 1080
 AAATTATCCA AACTTCTTCA ATGTTGGTGA TCAATGTATT TGGTTGCTTA CAGCTCCACG 1140
 CGGTGGATTC GTAAATTTGC AGTTTGTGTA ACAATTTCAA TTACAATGTG AAGATACGTG 1200
 TGATAAATCC TATGTAGAAG TGAAAGCTGA CGCTGATTTT CGACCTACTG GATATCGATT 1260
 10 TTGTTGTTTCG CGAGTGCCAC GTCATATTTT TCAATCTGCG ACAAACGAGA TGGTAGTAAT 1320
 ATTTTCGCGGT TTTGGTGGTG CGGGAAATGG CTTTAAAGCT AAAATTTGGT CAAACGTAGA 1380
 TGATGATATA GCTAATACAA TTGTAACAAC TGAAATGGCA AAAATTTTCGG AAAAAATACC 1440
 GAAGCTAACA GTTCCAATAG TTAAACTAT TACCACTCCT ACAATAACAA CTACTACTGC 1500
 TTTCATGATA TCACCCAAGA AAGGCAATGT CACCGCCACG AGAGTTGCTA TCACTACTAC 1560
 15 GCCGACTACT ACAATTACTA CGACTATTGC CGGTACGGTA CCAATCACCG TAACTAATAA 1620
 TACTACCCCT GTAGTAAGTG AAACCTTACC ATCATTGCCA GTCAAGATTC GAAACAAAAT 1680
 AGGTGCATGC GAATGTGGTG AATGGACAGA ATGGACAGGT CCATGCTCTC AAGAATGTGG 1740
 CGGTTGCGGA AAACGTCTTC GAACACGTCA GTGTTTCATCA GATACGGAAT GTAGAACAGA 1800
 AGAAAAACGT GCGTGTGCTT TTAAAGTTTG CCCATACGGG ACTAATTTCC TTATCAATAA 1860
 20 TGGAGAGTTT CATATACTTT GGAAGGGCTG CTGTGTTGGT CTATTCCGAT CGGGAGATAT 1920
 GTGTTTCAGCA CTTGATGATA ACGAGAATCC ATTTCTGAAA TTTCTAGAAT CACTGTTGAA 1980
 CATGCAAGAT TCTCGAAAAA ACGATAATTT GCCTGACTCG AAAAAAGAAGT GA 2032

(2) INFORMATION FOR SEQ ID NO:33:

25 (i) SEQUENCE CHARACTERISTICS:
 (A) LENGTH: 2028 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: single
 (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(ix) FEATURE:

(A) NAME/KEY: CDS

(B) LOCATION: 1..2028

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:33:

5 CAG TCA GTG GAA AAT TAT CGA ACG CAG AAA GCA TCA CGA AAT ACG TTA 48
Gln Ser Val Glu Asn Tyr Arg Thr Gln Lys Ala Ser Arg Asn Thr Leu
1 5 10 15

GAT CAC ATC AAA CAA CTT ATC ACC TTG AAC GTA CAA AGA GAG ATT GGA 96
Asp His Ile Lys Gln Leu Ile Thr Leu Asn Val Gln Arg Glu Ile Gly
10 20 25 30

AAC ATA GAT GAT AAG ACA TTA GCT GAT GAA ATA GTA TTA CAA CGA CGG 144
Asn Ile Asp Asp Lys Thr Leu Ala Asp Glu Ile Val Leu Gln Arg Arg
35 40 45

GAT CCT GAG GCA AAA TGG CAT CAT AAT GAA CTA TTC ATT AAT GAT CCA 192
Asp Pro Glu Ala Lys Trp His His Asn Glu Leu Phe Ile Asn Asp Pro
15 50 55 60

GAT GCA TAC TAT CAA GGC GAT GTC GAT TTG TCG GAA AAA CAA GCC GAA 240
Asp Ala Tyr Tyr Gln Gly Asp Val Asp Leu Ser Glu Lys Gln Ala Glu
65 70 75 80

ATT CTA AGC GAA CAT TTT AAA AAT GAA ATT GCT TTA ACA GAG AAA GAC 288
Ile Leu Ser Glu His Phe Lys Asn Glu Ile Ala Leu Thr Glu Lys Asp
20 85 90 95

GAC ACA ATA ATA CGG CGA AAA AAG AGC ATT GGT CGT GAA CCA TTT TAC 336
Asp Thr Ile Ile Arg Arg Lys Lys Ser Ile Gly Arg Glu Pro Phe Tyr
25 100 105 110

GTA AGA TGG AAT CAT AAA CGT CCC ATT AGC TAT GAA TTT GCG GAA AGT 384
Val Arg Trp Asn His Lys Arg Pro Ile Ser Tyr Glu Phe Ala Glu Ser
115 120 125

ATT CCA TTA GAA ACA CGT AGA AAA ATT CGT TCA GCA ATA GCA ATG TGG 432
Ile Pro Leu Glu Thr Arg Arg Lys Ile Arg Ser Ala Ile Ala Met Trp
30 130 135 140

GAA GAA CGA ACA TGC ATA CGA TTC CAA GAA AAT GGC CCA AAC GTA GAT 480
Glu Glu Arg Thr Cys Ile Arg Phe Gln Glu Asn Gly Pro Asn Val Asp
145 150 155 160

CGA ATT GTA TTT AAC GAC GGT GGC GGT TGT TCA AGT TTT GTC GGC CGA 528
Arg Ile Val Phe Asn Asp Gly Gly Gly Cys Ser Ser Phe Val Gly Arg
35 165 170 175

ACA GGA GGC ACG CCA GGA ATT TCA ATT TCA ACA CCA GGA TGT GAT ATT 576
Thr Gly Gly Thr Pro Gly Ile Ser Ile Ser Thr Pro Gly Cys Asp Ile
40 180 185 190

159

5 TCA CTG TTG AAC ATG CAA GAT TCT CGA AAA AAC GAT AAT TTG CCT GAC 2016
Ser Leu Leu Asn Met Gln Asp Ser Arg Lys Asn Asp Asn Leu Pro Asp
660 665 670

TCG AAA AAG AAG 2028
Ser Lys Lys Lys
675

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 676 amino acids
- (B) TYPE: amino acid
- (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:34:

160

Arg Ile Val Phe Asn Asp Gly Gly Gly Cys Ser Ser Phe Val Gly Arg
 165 170 175
 Thr Gly Gly Thr Pro Gly Ile Ser Ile Ser Thr Pro Gly Cys Asp Ile
 180 185 190
 5 Ile Gly Ile Ile Ser His Glu Ile Gly His Thr Leu Gly Ile Phe His
 195 200 205
 Glu Gln Ala Arg Arg Asp Gln Lys Asn His Ile Phe Ile Asn Tyr Asn
 210 215 220
 10 Asn Ile Pro Ser Ser Arg Trp Asn Asn Phe Phe Pro Leu Ser Glu Tyr
 225 230 235 240
 Glu Ala Asp Met Phe Asn Leu Pro Tyr Asp Thr Gly Ser Val Met His
 245 250 255
 Tyr Gly Ser Tyr Gly Phe Ala Arg Asn Pro Tyr Glu Pro Thr Ile Thr
 260 265 270
 15 Thr Arg Asp Lys Phe Gln Gln Tyr Thr Ile Gly Gln Arg Glu Gly Pro
 275 280 285
 Ser Phe Leu Asp Tyr Ala Ser Ile Asn Ser Ala Tyr Arg Cys Thr Glu
 290 295 300
 20 Gln Cys Ala Asp Met His Cys Asp His Asn Gly Tyr Pro Asp Pro Asn
 305 310 315 320
 Asn Cys Ala Lys Cys Leu Cys Pro Asp Gly Phe Ala Gly Arg Thr Cys
 325 330 335
 Gln Phe Val Gln Tyr Thr Ser Cys Gly Ala Leu Ile Lys Ala Arg Lys
 340 345 350
 25 Met Pro Val Thr Ile Ser Ser Pro Asn Tyr Pro Asn Phe Phe Asn Val
 355 360 365
 Gly Asp Gln Cys Ile Trp Leu Leu Thr Ala Pro Arg Gly Gly Phe Val
 370 375 380
 30 Asn Leu Gln Phe Val Glu Gln Phe Gln Leu Gln Cys Glu Asp Thr Cys
 385 390 395 400
 Asp Lys Ser Tyr Val Glu Val Lys Ala Asp Ala Asp Phe Arg Pro Thr
 405 410 415
 Gly Tyr Arg Phe Cys Cys Ser Arg Val Pro Arg His Ile Phe Gln Ser
 420 425 430
 35 Ala Thr Asn Glu Met Val Val Ile Phe Arg Gly Phe Gly Gly Ala Gly
 435 440 445

09003574-040699

35

(i) SEQUENCE CHARACTERISTICS:
 (A) LENGTH: 31 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: single
 (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

31

5

(ii) MOLECULE TYPE: DNA (genomic)

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:36:

31

124